NO 9/6

STIC-ILL

310243

From: Goldberg, Jeanine
Sent: Wednesday, September 06, 2000 8:14 AM
To: STIC-ILL
Subject: hGT1 gene

Please pull

1. E. J. of Human Genetics, Vol 6, pg 89-94, 1998
2. AMERICAN JOURNAL OF MEDICAL GENETICS, (1999 Dec 15) 88 (6) 694-9.
3. MOLECULAR PSYCHIATRY, (1999 Jan) 4 (1) 53-7. Journal code: CUM. ISSN: 1359-4184.

THANK YOU

Jeanine Enewold Goldberg
1655
CM1--12D11
306-5817

LC
9/8
VDS
NOS

COMPLETED

European Journal of Human Genetics (1998) 6, 89–94

ORIGINAL PAPER

The characterization and sequence analysis of thirty CTG-repeat containing genomic cosmid clones

Robert A Philibert[1], Nina Horelli-Kuitunen[2], Adelaide S Robb[1], Yu-Hsien Lee[1], Robert T Long[1], Patricia Damschroder-Williams[1], Brian M Martin[1], Miles B Brennan[1], Aarno Palotie[2] and Edward I Ginns[1]

[1]*Clinical Neuroscience Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892-4405, USA*
[2]*Department of Clinical Chemistry, Laboratory of Molecular Genetics, Helsinki University Central Hospital, Helsinki, Finland*

We have systematically isolated and characterized DNA containing large CTG ($n > 7$) repeats from a human cosmid genomic DNA library. Using a CTG_{10} probe, more than 100 cosmid clones were identified, and 30 of these have been extensively characterized. The sequenced cosmids contain repeats that are between three and 19 perfect units (average 10 perfect repeats). The cosmids map to at least 12 different chromosomes. Sequence analysis of flanking regions suggests that more than one third of the repeats occur in exons, and many share strong sequence identity with databank sequences, including the gene involved in dentatorubral pallidoluysian atrophy (DRPLA). Genotyping of human DNA samples demonstrates that more than half of the repeats are polymorphic. This and similar collections of clones containing trinucleotide repeats should aid in the identification of genes that may contain expansions of trinucleotide repeats involved in human disease.

Keywords: trinucleotide repeat; cosmid; fluorescent *in situ* hybridization (FISH); sequence analysis

Introduction

Pathological expansion of trinucleotide repeats[1] is responsible for several human diseases, including Huntington's chorea,[2] spinocerebellar ataxias,[3] spinal bulbar muscular atrophy[4] and dentatorubral pallidoluysian atrophy (DRPLA)[5] and myotonic dystrophy.[6,7] In each of these disorders, normally polymorphic repetitive DNA regions of between 10 and 30 perfect CTG or CAG units expand to greater than 40 units, resulting in disruptions of gene function.

Previously, investigators have attempted to isolate and characterize segments of DNA containing large CTG repeats from cDNA rather than genomic libraries.[8–11] While all CTG repeats associated with disease[1] should be represented, their isolation using cDNA libraries can be quite difficult for several reasons. First, low copy number, unstable, or tissue specific RNAs may be under-represented or completely absent from certain cDNA libraries.[12] Second, this approach of screening for cDNA will not identify trinucleotide containing regions in introns or in regions flanking genes. Third, trinucleotide repeats may not subclone in

Correspondence: Robert A Philibert, Clinical Neuroscience Branch, National Institute of Mental Health, Building 49, Room B1EE16, 49 Convent Drive, MSC 4405, Bethesda, MD 20892, USA
Accepted 10 October 1997

the vectors commonly used to generate cDNA libraries.

In order to circumvent these difficulties and to obtain as complete a representation as possible, we isolated and characterized trinucleotide repeats from a genomic DNA library and directly sequenced the cosmid genomic DNA inserts.[13] In this report, we describe our findings from the study of 30 trinucleotide repeat containing genomic clones.

Materials and Methods

The cosmid library was constructed from human DNA, partially digested with Sau 3A, ligated into the SuperCos I cosmid vector (Stratagene, La Jolla, CA), and packaged using the Gigapak II (Stratagene, La Jolla, CA). Positive colonies were identified by hybridization using the oligonucleotide probe CTG_{10}.[14] Briefly, three replicate cosmid colony lifts were prepared using ICN (Costa Mesa, CA) Biotrans membrane and then prehybridized for one hour in buffer (5 × SSPE, pH 7.0, 10 × Denhardt's solution, 0.05% SDS, and 10 µg/ml sheared *E. coli* DNA), and finally hybridized overnight in buffer (5 × SSPE, pH 7.0, 5 × Denhardt's solution, 0.1% SDS, and 10 µg/ml sheared *E. coli* DNA) containing 5' P^{32} labeled CTG_{10} probe. Individual sets of filters were then washed initially with two consecutive 5 min washes that were followed by more stringent washes at either 60°C, 70°C or 80°C for 15 min in 6 × SSPE. Filters were then exposed to X-ray film (Kodak X-OMAT-AR) for approximately 16 h at –70°C. Cosmid DNA was prepared and sequenced using either manual radioactive or automated fluorescent methods as described previously.[13]

PCR Amplification

PCR amplification of trinucleotide repeat containing DNA was performed using standard PCR buffer (10 mM Tris-HCl (pH 8.3), 50 mM KCl, 0.001% gelatin, 2 mM $MgCl_2$, 200 µM of each deoxynucleotide), 0.8 µM primers and 10% DMSO. Taq polymerase and genomic DNA concentrations were 2.5 U/100 µl and 50 ng/100 µl, respectively. The thermal cycling parameters for amplification were: initial denaturation of 95°C for 5 min, then 45 cycles of 95°C × 1 min, 65°C × 30 s, and 72°C × 2 min, followed by an final extension at 72°C × 10 min.

Polymorphism Analysis

Polymorphism analysis was conducted using DNAs from more than 30 unrelated individuals. The PCR products were exposed to electrophoresis at 1700 volts for 2–3 h on a 6% denaturing polyacrylamide sequencing gel. The separated PCR products were then electroblotted on to a Hy-bond N+ membrane (Amersham UK), hybridized overnight at 42°C in buffer (0.25 M NaCl, 0.125 M $NaPO_4$, 10% polyethylene glycol (MW6000), and 6% SDS) to a ^{32}P-labeled CTG_{10} probe, then washed, first at room temperature and then at 37°C, for 1 h with wash buffer (2 × SSC/1% SDS). Filters were then exposed overnight to Kodak X-OMAT-AR film at –70°C. The size of PCR products was determined by comparison with DNA sequencing ladder DNA fragments.

Sequence Analysis

Sequence analysis was performed using the BLAST[15] and GRAIL (Oakridge National Laboratory)[16,17] programs. Database comparisons and analyses were conducted on the cosmid DNA sequences with and without the trinucleotide repeat regions (see Results).

Chromosome Localization and Subchromosome Localization

Chromosomal assignment of the trinucleotide repeats was performed by PCR of somatic cell hybrid DNA (MPD-5000) from Bios Laboratories (New Haven, Connecticut).

Target Material for Fluorescence *in situ* Hybridization (FISH)

Peripheral blood lymphocytes were cultured according to standard protocols, and cells were treated with 5-bromodeoxyuridine (Brdu) at early replicating phase to induce banding pattern.[18] Slides were stained with Hoechst 33258 (1 µg/ml) for 10 min and exposed to UV light (302 nm) for 30 min.[19] Before hybridization metaphase slides were pretreated with RNAse (100 µg/ml) and pepsin (20 µg/ml).

Probes for FISH

CTG-containing cosmids were labeled with biotin 11-dUTP (Sigma Chemicals) by nick translation according to standard protocols (Nick Translation Kit, BRL).

FISH

The FISH procedure was carried out using 50% formamide, 10% dextran sulfate in 2 × SSC as described earlier.[19–22] Repetitive sequences were suppressed with 10–30 fold excess of COT-1 DNA (BRL, Gaithersburg, MD). After overnight incubation, nonspecific hybridization signals were eliminated by washing the slides with 50% formamide/2 × SSC, twice with 2 × SSC, and once with 0.5 × SSC at 45°C. Specific hybridization signals were visualized using FITC-conjugated Avidin (Vector Laboratories) and slides were counterstained with DAPI (4'–6'-diamino-2-phenylindole)(0.025 µg/ml). Only double spot signals were considered to be specific hybridizations. A multi-color image analysis was used for acquisition, display and quantification of hybridization signals of metaphase chromosomes. The system consists of a Photometrics PXL camera (Photometrics Inc, Tucson, AZ) attached to a PowerMac7100/Av workstation. IPLab software controls the camera operation, image acquisition and Ludl wheel.[23]

Results

From 800 000 human genomic cosmid clones screened with a ^{32}P-labeled CTG_{10} probe, 100 cosmids with positive hybridization signals were purified, and 30 were sequenced using the degenerate primer method.[13] Of these, 22 repeat sequences were unique whereas eight were represented twice. The chromosomal localization, length of the trinucleotide repeat, the heterozygosity, as well as the PCR primer sequences used to amplify the repeat region are shown in Table 1. Although the repeats average almost 10 perfect repeat

Table 1

Cosmid	Primer sequence	Genbank acc. no.	Chromosome[a] pcr	Chromosome[a] fish	No. rep.[b]	Het.[c]	Orf(1)[d]	Orf(1A)[e]	(Len)[f]	Prob.	Match acc. no.	Description
1	GGTGGTCACAGACTGGTACCAGTCC AGAAGAGGAGAAATTAAGTCAGCTTCA	AFO21105	19	ND	12	84%	No	No	221			
11	CTCACGTTCGAGGACTCTGGGAA TCCCCGTTCTTGTCGCCTTCCTCC	AFO21106	4,7	1q32	8	0%	Excel	Good	248			
12	GTCCAGCCGGGGCAAGAAGGTG TGTGCGTGCCCCCCCGGCGGT	AFO21107	4	4q13	9	0%	No	No	188			
13	AAGAGAACCCTCGGCTGGCTCCGG TTTCTCAGCCCAGCTGCAGCTCC	AFO21108	22	ND	8	0%	No	No	235			
14	GAGCCCTGAGTGTGGCCCTTGGT TTGGGGTTCCTCCTGAGTCCTGAA	AFO21109	10	10q26	10	0%	No	No	304			
15	CACATCACTGCACTCCAGCCTAGGT GTCAAAGCCTGACCATGTAAAGG	AFO21110	1	1p31–32 17q11.2–21	9	0%	No	No	305	10^{-20}	GB[L30192] GB[L17786] GB[L17917]	Human STS Human STS Human STS
17	GGACTGGACTGAGGCTAGACTCTC CCCAACAACACTGGACACCCGGC	AFO21111	16	12p12–13	13	34%	Marg	Good	274			
18	AGGTGTGCTGGCCCTGGTGGGTAG CCCCTTCCCTCCCTCTACTGG	AFO21112	12	ND	16	68%	No	No	281	ND	EMB[D31840]	DRPLA cDNA
21	GGGCAACGGATTGACTACGCGTCC GCGAGAGCAGCCGCGGCCAGCTCA	AFO21113	20	ND	4	0%	No	No	167			
22	GGAAGGGTCTCTGGGTCTTTGTGG CCGAACGGCTCTTGGATGCCCATG	AFO21114	ND	1q42	3	0%	No	No	251	10^{-60}	EMB[X00264]	Human beta-LH gene
23	ACAGCACCGACTGCCCAGCCCCAT GCTGATACTTGGCGAGGTTTTGGT	AFO21115	17	10q22 17q11.2	14	45%	Excel	Excel	595	10^{-131}	EMB[D29801] CHLC GCT 10C02	Mouse ORF
27	GGGAAGAGTCCATGGAGGT AGGTTGTGATCGCTGAGGTACCTG	AFO21116	ND	ND	9	5%	Excel	Excel	770			
34	GGCTAAGGTTGTGTGGAAGGCGCT GAAGAGGACGATGACGAGGAAGAG	AFO21117	12	ND	5	0%	Good	Excel	425			
37	AGAAGGAGGGCAGGGTGAGCCAG AAGACATCCTCTTGAAGAGGGTGG	AFO21118	17	17q11.2	5	0%	Excel	Excel	495	10^{-82}	EMB[D29801]	Mouse ORF
47	GATGGCTTGGGGTGGTGGTTATGA GCAGCAGGAGGAAGAGGAAGT	AFO21119	7	ND	13	0%	Marg	No	710	10^{-16}	GB[118120]	Human STS
56	ATTGTTTCCTGAAAGACCACACTA CCTCATCCCCAGTTCTGCTCACTGT	AFO21120	15	15q14–15	11	0%	No	No	240	10^{-40}	EMB[X76569] CHLC GCT 6B06	Human clone wgla4
62	CGGGCAGTTAGCACCCAGGTGGG GAAAAGCCCCAAGGACCAAGTCAG	AFO21121	17	17p11.2–12	10	0%	No	No	184			
66	CTGCAGCATTAGCAGGGCTCTGGC CTGCCACATGAGCTTCCTCTGTTG	AFO21122	22	22q11.2	13	50%	Marg	No	565		CHLC GCT 10C10	
67	TTGCAGCAGGCCGAGGAGGAAGAA ACCTCACCCACGCCATGCCTCACC	AFO21123	6	6p21.3	11	89%	No	No	231			
68	CTGACTACACTCGGTGCCCAGCCT TCAAGACGCAGCACGAGGAGTTT	AFO21124	19	19p13.1–13.2	6	2%	Excel	Excel	574			
70	AGAGGAGGTGGCTCCTGCCCCAGT CCATGCCTCACCTCTGGGTCTGAC	AFO21125	6	6p21.3	10	82%	Marg	No	517		CHLC GCT 4B05	
71	AGTGTCTCTGAAACCCACCACCTA GAGAAGTTCTGGTCAACCTTGCC	AFO21126	17	17p13	11	80%	Marg	No	1011			
73	AATACAGCCAGGCCCCAAGATGGA AGTGCAGCAGCCCAGGGTGGACGG	AFO21127	1	1q2122	19	55%	Good	Marg	478			

Table 1 *Continued*

Cosmid	Primer sequence	Genbank acc. no.	Chromosome[a] pcr	Chromosome[a] fish	No. rep.[b]	Het.[c]	Orf(1)[d]	Orf(1A)[e]	(Len)[f]	Prob.	Match acc. no.	Description
74	CTGCAGGGGACAGCAGGCTGCCTT CCCAGGTGCAGAGATTCACCTGTT	AFO21128	12	12q24.2–24.3 16p11.2	12	43%	Excel	Excel	467	10^{-17}	GB[M10668]	Mouse DNA with homology to EBVIR3
86	AGCCCGCTGCTCCATCCCCAAGCC ACACCAGGTGGGCGTCTGCGCTGG	AFO21129	1	ND	9	44%	Marg	No	636	10^{-30}	GC[L10375]	Human CTG-10cDNA sequence
91	TTCCCGTGCATACAAGCCAGCCTGG TCTCCTCCTCCAGGAGCCTGG	AFO21130	1	ND	10	45%	Marg	No	165			
94	AGCCGCGTACCATCTACCTGTGCC AGATATGCATAGAGACGTGCAGGC	AFO21131	10	ND	9	32%	No	No	240			
99	AGGAGATGCCAGAGCCTCTGCTGG CTCCAGCCTGGGAGACACAGCAA	AFO21132	8	ND	12	64%	No	No	159			
102	GAGACTCTGTGCTGACGGTCCGCC CAGGCCCTGGCAGTAGGCACCGTG	AFO21133	7	ND	6	0%	Marg	Marg	1918			
104	AACCTCTCCTCCCTCAACGGAGTG TGCTTGGTTTCTAAATGCAGCTAG	AFO21134	11	ND	4	0%	No	No	156			

[a]Chromosomal localization as established by PCR of somatic cell hybrid panels or FISH methodology.
[b]Number of repeats.
[c]Heterozygosity (HET) represents the percentage of individuals that have two alleles of different sizes.
[d]ORF(1) and [e]ORF(1A) denote the GRAIL subroutine used to analyze the sequence. For GRAIL 1 (ORF(1), 'Excellent' (EXCEL), 'Good', and 'Marginal' (MARG) denote probabilities of approximately 100%, 70% and <50% respectively, of the trinucleotide repeat being contained within an open reading frame. Before submission to GRAIL and BLAST for analyses, the repetitive CTG or CAG motif was removed in frame from the sequence.
[f]LEN represents the length of sequence in base pairs prior to removal of the trinucleotide motif submitted to GRAIL for analysis. Sequence identity comparison and probability calculations (PROB) were performed using BLAST.[18] Sequences with a greater than 10^{-20} match probability are reported under the column heading MATCH ACC. NO. ND=Not Determined.

units, some have additional short repeats adjacent to the CTG repeats. For example, CTG-1 and CTG-15, have large CAA or GAA repeats adjacent to the CTG repeat. The probability of each trinucleotide repeat being located within an exon was determined by the Gene Recognition and Analysis Interlink (GRAIL) Program[16] (Table 1). In order to avoid difficulties inherent in the analysis of repetitive DNA regions, the trinucleotide repeat was deleted before the GRAIL analyses. Despite the short length of many of the sequences submitted for GRAIL analysis, approximately one third of the sequences had a good or excellent probability of occurring in exons.

These trinucleotide-depleted sequences were also submitted using the Basic Local Alignment Search Tool (BLAST)[15] for comparison to Genbank and Swiss Prot data banks. DNA in nine clones showed at least a mild degree ($p < 10^{-15}$) of sequence homology to database entries. Regions of cosmids CTG-37 and CTG-23 show almost complete sequence identity to a mouse open reading frame (ORF) encoding a central nervous system protein, while CTG-22 shows strong sequence identity with a region of beta-luteinizing hormone. Cosmid CTG-56 shows considerable sequence identity with wg1A (EMB(X76569)), a previously isolated trinucleotide repeat.[24] Cosmid CTG-18 contains the genomic clone of the DRPLA cDNA clone.[5] CTG-86 is similar to CTG-B10, a trinucleotide-containing clone previously isolated from a human brain cDNA library.[9] For the other 21 CTG repeat-containing clones, including five with a good or excellent probability of occurring in exons, no sequence homology to database entries was identified.

Discussion

Our findings suggest that the direct sequencing of genomic trinucleotide repeat-containing clones is useful for studying the involvement of these repetitive regions in human disease. With a few exceptions,[24–27] previous attempts to characterize large CTG repeats have utilized cDNA libraries,[8,9] resulting in a bias toward over-represented, more clonable, and/or more abundant transcripts. This makes the isolation of the interesting, rare or less stable cDNAs difficult, and is in contrast to procedures using genomic libraries which tend to have a less biased representation of the total candidate gene pool.

The direct sequencing of cosmid clones[13] has several advantages. First, the large trinucleotide repeats which tend to be eliminated using smaller plasmids are more stable in cosmids. Second, analysis of the genomic DNA sequence surrounding the repeat allows us to determine whether the repeat could be located within an exon. Third, the additional sequence available in a cosmid can be used to generate FISH probes, allowing for subchromosomal location of clone. The isolation of genomic trinucleotide repeats by subcloning filter hybridization enriched, PCR amplified, Mbo-I digested genomic fragments can be an alternative to generation of a primary library,[24] but these repetitive regions are often difficult to amplify,[28] resulting in the isolation of smaller, less GC-rich repeats that provide much less sequence information.

Using an approach in which the repetitive CTG sequence is removed, GRAIL analyses indicated that at least one third of these sequences has good or excellent probability of being found in a coding exon. This may underestimate the frequency of ORFs since at least one sequence, CTG-18, which stands for part of the DRPLA locus, was not detected by this GRAIL analysis. This omission may have occurred because GRAIL sometimes fails to recognize coding exons less than 100 bp in length. In an analysis of genomic CTG repeat sequences obtained from GENBANK[29] Stallings concluded that one third of CTG repeats and almost all CAG repeats were located in exons. Our results are in good agreement with these previous findings.

Comparison of the repeat sequences in our study with those in GENBANK demonstrates that several have significant sequence identity with previously described DNA sequences. The finding that CTG-18 is a partial genomic clone for the DRPLA cDNA illustrates the usefulness of this approach to search for trinucleotide repeats that may be involved in human disease. Both CTG-23 and CTG-37 have considerable sequence identity with different parts of murine ORF (D29801). Interestingly, GRAIL predicts that, like the CAG repeats from the mouse ORF (D29801), the repeats from CTG-23 and CTG-37 are exonic in humans. However, the murine repeats are much smaller, being only 2 or 3 CAG units in length. This suggests that the trinucleotide repeats on chromosome 17 represented by CTG-23 and CTG-37 expanded after the divergence of human and mouse genomes.

With two exceptions, CTG-11 and CTG-17, the FISH data confirm the somatic cell PCR localization results. Two of the repeat-containing cosmids, CTG-74 and CTG-15 map by FISH to two distinct loci. This observation may result from the presence of multiple

copies of these trinucleotide repeats or suggest the presence of a gene family of related sequences. This is not surprising since at least one repeat, CTG-47 gives four allele fragments on PCR amplification of human genomic DNA. However, unlike CTG-74 and CTG-15, chromosome localization performed using somatic cell hybrids suggests that all the loci encoding CTG-47 repeat sequence are on chromosome 7.

In summary, we demonstrate that direct sequencing of cosmid clones from a genomic library is a useful approach to isolating and characterizing DNA sequences containing trinucleotide repeats that could be involved in human disease. The chromosomal and sub-chromosomal localization data presented here provide sequences that may help to identify candidate genes for diseases mapping nearby or in yet to be localized syndromes.

Acknowledgements

We would like to thank Ms Kay Kuhns and Ms Liz Alzona for manuscript preparation. RAP was in part supported by the Pharmacology Research Training Program, NIGMS.

References

1 Warren ST, Nelson DL: Advances in molecular analysis of fragile X syndrome. *JAMA* 1994; **271**: 536–542.
2 The Huntington's Disease Collaborative Research Group: A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 1993; **72**: 971–983.
3 Koshy BT, Zoghbi HY: The CAG/polyglutamine tract diseases: gene products and molecular pathogenesis. *Brain Pathol* 1997; **7**: 927–942.
4 LaSpada AR *et al*: Androgen receptor gene mutation in linked spinal and bulbar muscular atrophy. *Nature* 1992; **352**: 77–79.
5 Koide R *et al*: Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). *Nature Genet* 1994; **6**: 14–18.
6 Harley HG *et al*: Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 1992; **355**: 545–546.
7 Buxton J *et al*: Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* 1992; **355**: 547–548.
8 Riggins GJ *et al*: Human genes containing polymorphic trinucleotide repeats. *Nature Genet* 1992; **2**: 186–191.
9 Li S-H *et al*: Novel triplet repeat containing genes in human brain: cloning, expression and length polymorphisms. *Genomics* 1993; **16**: 572–579.
10 Neri C *et al*: Survey of CAG/CTG repeats in human cDNAs representing new genes: candidates for inherited neurological disorders. *Hum Mol Genet* 1996; **5**(7): 1001–1009.
11 Jiang JX, Deprez RH, Zwarthoff EC, Riegman PH: Characterization of four novel CAG repeat-containing cDNAs. *Genomics* 1995; **30**(1): 91–93.
12 Kimmel AR: Selection of clones from libraries: overview. In: Berger S, Kimmel AR (eds). *Methods in Enzymology*. Academic Press: New York, 1987, vol 152, pp 393–398.
13 Philibert RA *et al*: Direct sequencing of trinucleotide repeats from cosmid genomic DNA templates. *Anal Biochem* 1995; **225**: 372–374.
14 Wallace RB, Miyada CG: Guide to molecular cloning techniques. In Berger SL, Kimmel AR (eds). *Methods in Enzymology*. Academic Press, New York, 1987, vol 152, pp 432–442.
15 Altschul SF *et al*: Basic local alignment search tool. *J Mol Biol* 1990; **215**: 403–410.
16 Uberbacher EC, Mural RJ: Locating protein-coding regions in human DNA sequences by a multiple sensor-neural network approach. *Proc Nat Acad Sci USA* 1991; **88**: 11261–11265.
17 Shah MB *et al*: User's guide to GRAIL and GENQUEST (Sequence analysis, gene assembly and sequence comparison systems) E-mail servers and XGRAIL (version 1.2) and XGENQUEST (version 1.1) client-server systems. Available by anonymous ftp to arthur.epm.ornl.gov (128.219.9.76) from directory pub/xgrail or pub/xgenquest as file manual.grail-genquest. 1994.
18 Lemieux N, Dutrillaux B, Viegas-Pequignot E: A simple method for simultaneous R- or G-banding and fluorescence *in situ*hybridization of small single copy genes. *Cytogenetics Cell Genet* 1992; **59**: 311–312.
19 Tenhunen K *et al*: Molecular cloning, chromosomal assignment, and expression of the mouse aspartylglucosaminidase gene. *Genomics* 1995; **30**: 244–250.
20 Lichter P *et al*: Rapid detection of human chromosome 21 aberrations by *in situ* hybridization. *Proc Natl Acad Sci USA* 1988; **85**: 9664–9668.
21 Pinkel D *et al*: Fluorescence *in situ* hybridization with human chromosome-specific libraries: detection of trisomy 21 and translocations of chromosome 4. . *Proc Natl Acad Sci USA* 1988; **85**: 9138–9142.
22 Rytkonen EM *et al*: The human gene for xanthine dehydrogenase (XDH) is localized on chromosome band 2q22. *Cell Genet* 1995; **68**: 61–63.
23 Heiskanen M *et al*: Visual mapping by fiber-FISH. *Genomics* 1995; **30**: 31–36.
24 Armour JAL, Neumann R, Gobert S, Jeffreys AJ: Isolation of human simple repeat loci by hybridization selection. *Hum Mol Genet* 1994; **3**: 599–605.
25 Gastier JM *et al*: Survey of trinucleotide repeats in the human genome: assessment of their utility as genetic markers. *Hum Mol Genet* 1995; **4**: 1829–1836.
26 Gastier JM *et al*: Development of a screening set for new $(CAG/CTG)_n$ dynamic mutations. *Genomics* 1996; **32**: 75–85.
27 Sunden SLF *et al*: Chromosomal assignment of 2900 tri- and tetranucleotide repeat markers using NIGMS somatic cell hybrid panel 2. *Genomics* 1996; **32**: 15–20.
28 Reiss AJ *et al*: Frequency and stability of the fragile X mutation. *Hum Mol Genet* 1994; **3**: 393–398.
29 Stallings RL: Distribution of trinucleotide microsatellites in different categories of mammalian genomic sequence: implication for human genetic diseases. *Genomics* 1994; **21**: 116–121.